

# MSE1 Important Principles

## Definition 1.3 (One-to-One functions (Injective))

A function  $f : A \rightarrow B$  is called **one-to-one** (or **injective**) if different elements in  $A$  map to different elements in  $B$ . In other words, if  $f(a_1) = f(a_2)$ , then  $a_1 = a_2$ . This property ensures that no two distinct elements in  $A$  are mapped to the same element in  $B$ .



Graphically, a function is one-to-one if no horizontal line intersects the graph of the function at more than one point.

## Definition 1.4 (Onto Functions (Surjective))

A function  $f : A \rightarrow B$  is called **onto** (or **surjective**) if every element in  $B$  is the image of at least one element in  $A$ . In other words, for every  $b \in B$ , there exists at least one  $a \in A$  such that  $f(a) = b$ . This property ensures that the function “covers” the entire set  $B$ .



## Definition 1.5 (Inverse Functions)

Let  $f$  be a one-to-one correspondence from the set  $A$  to the set  $B$ . The inverse function of  $f$  is the function that assigns to an element  $b$  belonging to  $B$  the unique element  $a$  in  $A$  such that  $f(a) = b$ . The inverse function of  $f$  is denoted by  $f^{-1}$ . Hence,  $f^{-1}(b) = a$  when  $f(a) = b$ .



## Definition 1.6

Let  $f : B \rightarrow C$  and  $g : A \rightarrow B$  be two functions. The **composite function** of  $f$  and  $g$ , denoted by  $f \circ g$ , is a function from  $A$  to  $C$  defined by

$$(f \circ g)(x) = f(g(x)),$$

for every  $x \in A$ .



## Definition 1.8 (Base-10 Logarithms)

$$\log x = y \iff 10^y = x$$

*Verbally:*  $\log x$  is the exponent in the power of 10 that gives  $x$



### Properties of base-10 logarithms

- Log of a Product:

$$\log xy = \log x + \log y$$

*Verbally:* The log of a product equals the sum of the logs of the factors.

- Log of a Quotient:

$$\log \frac{x}{y} = \log x - \log y$$

*Verbally:* The log of a quotient equals the log of the numerator minus the log of the denominator.

- Log of a Power:

$$\log x^y = y \log x$$

*Verbally:* The log of a power equals the exponent times the log of the base.

### Definition 1.9 (Logarithm with Any Base)

*Algebraically:*

$$\log_b x = y \text{ if and only if } b^y = x, \quad \text{where } b > 0, b \neq 1, \text{ and } x > 0$$

*Verbally:*

$\log_b x = y$  means that  $y$  is the exponent of  $b$  that gives  $x$  as the answer.



### Definition 1.10 (Common Logarithm and Natural Logarithm)

*Common:* The symbol  $\log x$  means  $\log_{10} x$ .

*Natural:* The symbol  $\ln x$  means  $\log_e x$ , where  $e$  is a constant equal to 2.71828182845...



### The Change-of-Base Property of Logarithms

$$\log_a x = \frac{\log_b x}{\log_b a} \quad \text{or} \quad \log_a x = \frac{1}{\log_b a} (\log_b x)$$

### Properties of Logarithms

The Logarithm of a Power:

$$\log_b x^y = y \log_b x$$

*Verbally:* The logarithm of a power equals the product of the exponent and the logarithm of the base. The Logarithm of a Product:

$$\log_b(xy) = \log_b x + \log_b y$$

*Verbally:* The logarithm of a product equals the sum of the logarithms of the factors. The Logarithm of a Quotient:

$$\log_b \frac{x}{y} = \log_b x - \log_b y$$

*Verbally:* The logarithm of a quotient equals the logarithm of the numerator minus the logarithm of the denominator.

### Commutative, Associative, and Distributive Laws

- Commutativity:  $a + b = b + a$
- Associativity:  $(a + b) + c = a + (b + c)$
- Distributive property:  $a \times (b + c) = a \times b + a \times c$

### Closure Property

**Addition:** The set of integers is closed under the operation of addition. This means that if you take any two integers and add them together, the sum will always be an integer. For example, if  $a$  and  $b$  are integers, then  $a + b$  is also an integer. This closure property ensures that the set of integers is stable and complete under addition, meaning that no matter how many times you add integers together, the result will remain within the set of integers.

**Multiplication:** The set of integers is also closed under multiplication. If you multiply any two integers, the product will always be an integer. For instance, if  $a$  and  $b$  are integers, then  $a \times b$  is also an integer. This property guarantees that the operation of multiplication, like addition, does not produce results outside the set of integers, thereby preserving the integrity of the set under multiplication.

## Boolean Algebra

1. Complements (NOT) are evaluated first.
2. Boolean products (AND) are computed next.
3. Boolean sums (OR) are evaluated last.

## Logical Operators

The complement, Boolean sum, and Boolean product correspond to the logical operators  $\neg$ ,  $\vee$ , and  $\wedge$ , respectively. In this context, 0 corresponds to FALSE (F) and 1 corresponds to TRUE (T).

### Fundamental Boolean Properties

- **Double Complement:**

$$\overline{\overline{x}} = x$$

- **Idempotent Laws:**

$$x + x = x, \quad x \cdot x = x$$

- **Identity Laws:**

$$x + 0 = x, \quad x \cdot 1 = x$$

- **Domination Laws:**

$$x + 1 = 1, \quad x \cdot 0 = 0$$

- **Unit Property:**

$$\overline{x} + x = 1$$

- **Zero Property:**

$$\overline{x} \cdot x = 0$$

### Structural Boolean Laws

- **Commutative Laws:**

$$x + y = y + x, \quad x \cdot y = y \cdot x$$

- **Absorption Laws:**

$$x + (x \cdot y) = x, \quad x \cdot (x + y) = x, \quad \overline{x} \cdot \overline{y} + x = \overline{y} + x$$

- **De Morgan's Laws:**

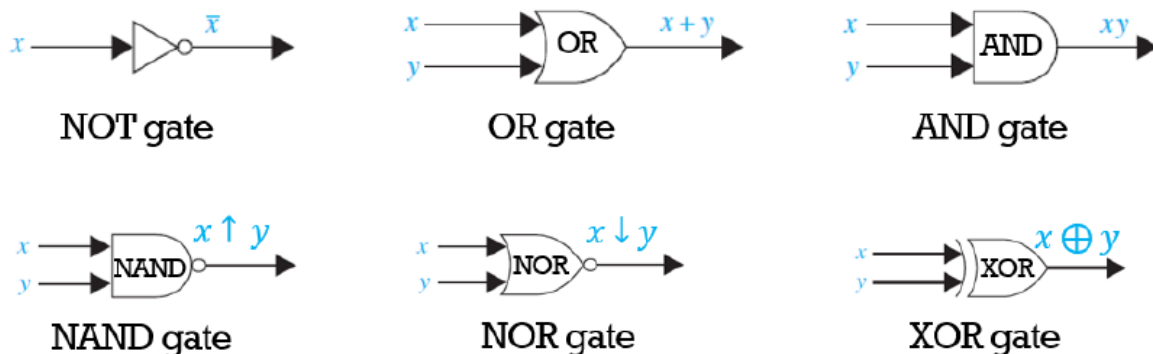
$$\overline{x \cdot y} = \overline{x} + \overline{y}, \quad \overline{x + y} = \overline{x} \cdot \overline{y}$$

- **Associative Laws:**

$$x + (y + z) = (x + y) + z, \quad x \cdot (y \cdot z) = (x \cdot y) \cdot z$$

- **Distributive Law:**

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z), \quad x + (y \cdot z) = (x + y) \cdot (x + z)$$



**Figure 4.4:** ANSI Symbols of Most Common Logic Gates. (Rosen, 2012).

# Probability

## Definition 5.2 (Sample Spaces, Outcomes, and Events)

**Sample Space:** The set of all possible outcomes of a random experiment is called the sample space, denoted by  $S$ . Outcomes can be discrete or continuous, depending on the nature of the experiment.

**Outcome:** A single possible result of a random experiment.

**Event:** Any subset of the sample space, which may consist of one or more outcomes.



- **Mutually Exclusive Events (or Disjoint):** Two events are mutually exclusive (or disjoint) if they cannot occur at the same time. Their intersection is empty:  $A \cap B = \emptyset$ . For example, the events of rolling a 2 and rolling a 4 on a die are mutually exclusive because they cannot happen simultaneously.
- **Collectively Exhaustive Events:** A set of events is collectively exhaustive if at least one of the events must occur. Together, they cover the entire sample space. For example, when rolling a die, the events  $A = \{\text{even numbers}\}$  and  $B = \{\text{odd numbers}\}$  are collectively exhaustive because every outcome is either even or odd.

Operation	Boolean Algebra	Logic	Set Theory
NOT	$\bar{x}$	$\neg x$	$A^c$ or $\bar{A}$ or $A'$
OR	$+$	$\vee$	$\cup$
AND	$\cdot$	$\wedge$	$\cap$
NAND	$\overline{x \cdot y}$	$\neg(x \wedge y)$	$(A \cap B)^c$ or $\overline{A \cap B}$
NOR	$\overline{x + y}$	$\neg(x \vee y)$	$(A \cup B)^c$ or $\overline{A \cup B}$
XOR (Symmetric Difference)	$x \oplus y$	$(x \wedge \neg y) \vee (\neg x \wedge y)$	$A \Delta B$ or $A \oplus B$
Difference	$x \cdot \bar{y}$	$x \wedge \neg y$	$A - B$ or $A \setminus B$

**Table 5.1:** Comparison of Operators in Boolean Algebra, Logic, and Set Theory

## Theorem 5.1 (Multiplication Rule)

Let an operation be described as a sequence of  $k$  steps. Assume the following conditions:

- There are  $n_1$  ways to complete step 1.
- There are  $n_2$  ways to complete step 2 for each way of completing step 1.
- There are  $n_3$  ways to complete step 3 for each way of completing step 2, and so on.

Then, the total number of ways to complete the entire operation is given by:

$$n_1 \times n_2 \times \cdots \times n_k.$$



#### Definition 5.4 (Permutation and Combination)

- **Permutation (Order Matters):** Different sequences are counted as distinct outcomes, leading to a higher count.
- **Combination (Order Does Not Matter):** Sequences are treated as identical, resulting in a lower count.



#### Definition 5.5 (Permutations of $n$ Objects)

The permutation of  $n$  objects, i.e. an ordered arrangement of  $n$  objects, is  $n!$  (read as "n factorial"), where:

$$n! = n \times (n - 1) \times (n - 2) \times \cdots \times 1$$



#### Definition 5.6 (Permutations of Subsets)

The number of permutations of subsets of  $r$  elements selected from a set of  $n$  different elements is

$$P_r^n = n \times (n - 1) \times (n - 2) \times \cdots \times (n - r + 1) = \frac{n!}{(n - r)!}$$



#### Definition 5.7 (Combinations)

The number of combinations of  $r$  elements selected from a set of  $n$  different elements is given by

$$C_r^n = \binom{n}{r} = \frac{n!}{r!(n - r)!}$$



#### Theorem 5.2

In algebra, the binomial coefficient is used to expand powers of binomials. According to the binomial theorem

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$$



### Theorem 5.3 (Rules of Probability)

- **Complement Rule:** The probability of the complement of event  $A$  is
$$P(\overline{A}) = 1 - P(A)$$
- **Empty Set Rule:** The probability of the empty set is 0, i.e.,
$$P(\emptyset) = 0$$
- **Addition Rule:** For any two events  $A$  and  $B$ , the probability of the union of events  $A$  and  $B$  is given by
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
- **Difference Rule:** The probability of the difference between events  $A$  and  $B$  is given by
$$P(A - B) = P(A) - P(A \cap B)$$
- **Subset Rule:** If  $A$  is a subset of  $B$  ( $A \subset B$ ), then
$$P(A) \leq P(B)$$



### Definition 6.1 (Conditional Probability)

For any two events  $A$  and  $B$  with  $P(B) > 0$ , the conditional probability of  $A$  given  $B$  is defined as:

$$P(A | B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0$$



### Axiom 6.1 (Axioms of Conditional Probability)

- **Axiom 1:** For any event  $A$ ,  $0 \leq P(A | B) \leq 1$ .
- **Axiom 2:** Conditional probability of  $B$  given  $B$  is 1, i.e.,  $P(B | B) = 1$ .
- **Axiom 3:** If  $A_1, A_2, A_3, \dots$  are disjoint events, then  $P(A_1 \cup A_2 \cup A_3 \dots | B) = P(A_1 | B) + P(A_2 | B) + P(A_3 | B) + \dots$



### Theorem 6.1 (Rules of Conditional Probability)

- **Complement Rule:** The probability of the complement of event  $A$  given  $C$  is
$$P(\overline{A} | C) = 1 - P(A | C)$$
- **Empty Set Rule:** The probability of the empty set given some event  $C$  is 0, i.e.,
$$P(\emptyset | C) = 0$$
- **Addition Rule:** For any two events  $A$  and  $B$ , the probability of the union of events  $A$  and  $B$  given another event  $C$  is
$$P(A \cup B | C) = P(A | C) + P(B | C) - P(A \cap B | C)$$
- **Difference Rule:** The probability of the difference between events  $A$  and  $B$  given another event  $C$  is
$$P(A - B | C) = P(A | C) - P(A \cap B | C)$$
- **Subset Rule:** If  $A \subset B$ , then
$$P(A | C) \leq P(B | C)$$



### Theorem 6.2 (Multiplication Rule)

For any two events  $A$  and  $B$ , the probability of both events occurring is given by:

$$P(A \cap B) = P(A)P(B | A) = P(B)P(A | B)$$

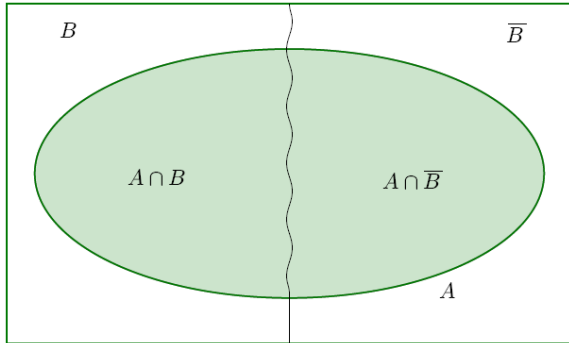


Figure 6.1:  $P(A) = P(A \cap B) + P(A \cap \bar{B})$

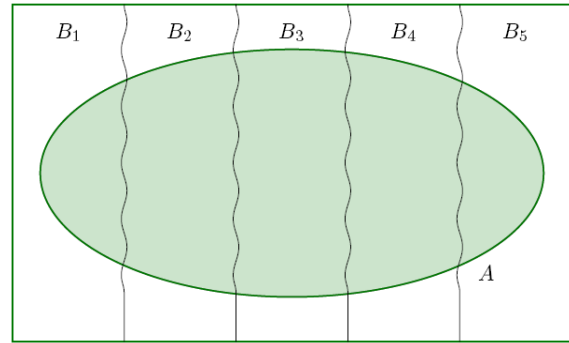


Figure 6.2:  $P(A) = \sum_{i=1}^5 P(A | B_i)P(B_i)$

### Definition 6.2 (Independence)

Two events are considered independent if the occurrence of one event does not affect the probability of the other event. In other words, the probability of one event does not depend on the occurrence of the other event. Two events  $A$  and  $B$  are independent if:

$$P(A | B) = P(A)$$

$$P(B | A) = P(B)$$

$$P(A \cap B) = P(A)P(B)$$



### Lemma 6.1

If  $A$  and  $B$  are independent then

- $A$  and  $\bar{B}$  are independent,
- $\bar{A}$  and  $B$  are independent,
- $\bar{A}$  and  $\bar{B}$  are independent.



### Theorem 6.4 (Independence and DeMorgan's Law)

If  $A_1, A_2, \dots, A_n$  are independent then

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - (1 - P(A_1))(1 - P(A_2)) \dots (1 - P(A_n))$$



**Warning!** A common misconception is to confuse independence with disjointness. However, these are fundamentally different concepts. Two events,  $A$  and  $B$ , are disjoint if the occurrence of one prevents the occurrence of the other, i.e.,  $A \cap B = \emptyset$ . In this case, knowing that  $A$  has occurred gives us complete information about  $B$  namely, that  $B$  cannot occur. This dependence means that disjoint events cannot be independent.



Concept	Description	Key Formulas
Disjoint	Events $A$ and $B$ cannot occur at the same time	$A \cap B = \emptyset$ $P(A \cup B) = P(A) + P(B)$
Independent	Occurrence of $B$ gives no information about $A$	$P(A   B) = P(A)$ $P(B   A) = P(B)$ $P(A \cap B) = P(A) \cdot P(B)$

**Table 6.1:** Comparison of Disjoint and Independent Events.

### Definition 6.3 (Independence of Multiple Events)

A set of events  $A_1, A_2, \dots, A_k$  are considered independent if the joint probability is equal to the product of the individual probabilities:

$$P(A_1 \cap A_2 \cap \dots \cap A_k) = P(A_1) \cdot P(A_2) \cdot \dots \cdot P(A_k)$$



### Theorem 6.5 (Bayes' Theorem)

For any two events  $A$  and  $B$ , where  $P(A) \neq 0$ , we have

$$P(B | A) = \frac{P(A | B) \cdot P(B)}{P(A)} = \frac{P(A | B) \cdot P(B)}{P(A | B) \cdot P(B) + P(A | \bar{B}) \cdot P(\bar{B})}$$

If  $B_1, B_2, B_3, \dots$  form a partition of the sample space  $S$ , and  $A$  is any event with  $P(A) \neq 0$ , we have

$$P(B_j | A) = \frac{P(A | B_j) \cdot P(B_j)}{\sum_i P(A | B_i) \cdot P(B_i)}$$



## Linear Algebra

### Definition 7.1 (Linear Equations)

A linear equation in the variables  $x_1, x_2, \dots, x_n$  is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b$$

where  $b$  and the coefficients  $a_1, \dots, a_n$  are constants.



### Definition 7.3 (Solutions to a System of Linear Equations)

A **solution** of a linear system in the variables  $x_1, x_2, \dots, x_n$  is a list of numbers  $(s_1, s_2, \dots, s_n)$  that satisfies all equations of the system when substituted for the variables  $x_1, x_2, \dots, x_n$ , respectively.

The set of all possible solutions is called its **solution set**. Two linear systems are called **equivalent** if they have the same solution set.



A system of linear equations has one of the following outcomes:

- (i) **No solutions**, when the equations are inconsistent.
- (ii) **Exactly one solution**, when there is a unique set of values satisfying all equations.
- (iii) **Infinitely many solutions**, when there are multiple sets of values that satisfy the equations.

We say a system is **consistent** if it has either one or infinitely many solutions. We say a system is **inconsistent** if it has no solution. We formalise this later in this chapter.

#### Definition 7.4

A **matrix** is a rectangular array of numbers arranged in rows and columns. The **coefficient matrix** of a linear system contains only the coefficients of the variables, while the **augmented matrix** includes an additional column for the constants from the right-hand side of the equations.



The three types of elementary row operations are:

- **Replacement:** Replace one row by the sum of itself and a multiple of another row.
- **Interchange:** Swap two rows.
- **Scaling:** Multiply all entries in a row by a nonzero constant.

Two matrices are called **row equivalent** if one can be transformed into the other through a sequence of elementary row operations.

#### Definition 7.6 (Echelon Forms)

A matrix is in **echelon form** if it satisfies the following conditions:

1. All zero rows are at the bottom.
2. Each leading entry of a row is in a column to the right of the leading entry of the row above it.
3. All entries in a column below a leading entry are zeros.

In addition to echelon form, a matrix may also be in **reduced row echelon form** (RREF), which has the following properties:

4. The leading entry in each nonzero row is 1.
5. Each leading 1 is the only nonzero entry in its column.



#### Theorem 7.1 (Existence Theorem)

A linear system is consistent if and only if an echelon form of the augmented matrix has no row of the form

$$\begin{bmatrix} 0 & \dots & 0 & b \end{bmatrix}$$

where  $b$  is nonzero.



#### Theorem 7.2 (Uniqueness of Reduced Echelon Form)

Each matrix is row equivalent to one and only one reduced echelon matrix



### The Row Reduction Algorithm

Here we describe an algorithm for turning any matrix into an equivalent (reduced) echelon matrix. This algorithm is the foundation of solving systems of linear equations using matrices.

1. Begin with the leftmost nonzero column. This is a pivot column, with the pivot position at the top.
2. Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry into the pivot position.
3. Use row replacement operations to create zeros in all positions below the pivot.
4. Apply steps 1-3 to the submatrix of all entries below and to the right of the pivot position. Repeat this process until there are no more nonzero rows to modify. (At this point we have reached an echelon form of the matrix.)
5. Beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot using row operations. If a pivot is not 1, make it 1 by a scaling operation. (This step produces the reduced echelon form of the matrix.)

### Theorem 7.3 (Uniqueness Theorem)

If a linear system is consistent, then the solution set contains either

- (i) a unique solution, when there are no free variables, or
- (ii) infinitely many solutions, when there is at least one free variable.



In linear algebra, vectors are often represented as **column matrices**. For example, the vector  $\mathbf{v} = (x_1, x_2)$  can be written as a column matrix:

$$\mathbf{v} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

We say that two vectors are **equal** if and only if their corresponding entries are equal.

### Definition 8.1

For each positive integer  $n$ , we let  $\mathbb{R}^n$  denote the collection of ordered  $n$ -tuples with each entry in  $\mathbb{R}$ . We often write these elements as  $n \times 1$  matrices. We define addition and scalar multiplication of vectors in  $\mathbb{R}^n$  in the same way as we do for  $\mathbb{R}^2$ . That is, we go coordinate-by-coordinate.



### Algebraic Properties of $\mathbb{R}^n$

Let  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$  and  $c, d \in \mathbb{R}$ . Then the following properties hold:

1. **Commutative Property of Addition:**  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ .
2. **Associative Property of Addition:**  $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ .
3. **Additive Identity:** There exists a vector  $\mathbf{0} \in \mathbb{R}^n$  such that  $\mathbf{u} + \mathbf{0} = \mathbf{u}$  for all  $\mathbf{u} \in \mathbb{R}^n$ .
4. **Additive Inverse:** For each  $\mathbf{u} \in \mathbb{R}^n$ , there exists a vector  $-\mathbf{u} \in \mathbb{R}^n$  such that  $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ .
5. **Distributive Property:**  $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$  and  $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ .
6. **Associative Property of Scalar Multiplication:**  $c(d\mathbf{u}) = (cd)\mathbf{u}$ .
7. **Multiplicative Identity:**  $1\mathbf{u} = \mathbf{u}$ .

### Definition 8.2 (Linear Combinations)

Given a set of vectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p \in \mathbb{R}^n$  and scalars  $c_1, c_2, \dots, c_p \in \mathbb{R}$ , the vector  $\mathbf{y}$  given by

$$\mathbf{y} = c_1\mathbf{v}_1 + \dots + c_p\mathbf{v}_p$$

is called a linear combination of  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$  with weights  $c_1, c_2, \dots, c_p$ .



### Using Matrices to Determine Linear Combinations

A vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n = \mathbf{b}$$

has the same solution set as the linear system whose augmented matrix is

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n & \mathbf{b} \end{bmatrix}.$$

More specifically, if the  $\mathbf{a}_i$ 's are in  $\mathbb{R}^m$  with

$$\mathbf{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, \quad \mathbf{a}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

then you would row reduce the matrix

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix}$$

$$\begin{matrix} \uparrow & \uparrow & \dots & \uparrow & \uparrow \\ \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n & \mathbf{b} \end{matrix}$$

to determine if there is some set of weights  $x_1, \dots, x_n$  that work.

### Definition 8.3 (Span of a Set of Vectors)

If  $\mathbf{v}_1, \dots, \mathbf{v}_p$  are in  $\mathbb{R}^n$ , then the set of all linear combinations of  $\mathbf{v}_1, \dots, \mathbf{v}_p$  is denoted by  $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$  and is called the subset of  $\mathbb{R}^n$  spanned by  $\mathbf{v}_1, \dots, \mathbf{v}_p$ . In other words, the span of  $\mathbf{v}_1, \dots, \mathbf{v}_p$  is all vectors that can be written in the form

$$c_1\mathbf{v}_1 + \dots + c_p\mathbf{v}_p$$

with  $c_1, \dots, c_p$  scalars.



**Definition 8.4 (Matrix Equation)**

If  $A$  is an  $m \times n$  matrix with columns  $\mathbf{a}_1, \dots, \mathbf{a}_n$ , and if  $\mathbf{x} \in \mathbb{R}^n$ , then the product of  $A$  and  $\mathbf{x}$ , denoted by  $A\mathbf{x}$ , is

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n$$

**Theorem 8.1**

If  $A$  is an  $m \times n$  matrix, with columns  $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$  and if  $\mathbf{b} \in \mathbb{R}^m$ , the matrix equation

$$A\mathbf{x} = \mathbf{b}$$

has the same solution set as the vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n = \mathbf{b}$$

which, in turn, has the same solution set as the system of linear equations with augmented matrix

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n & \mathbf{b} \end{bmatrix}.$$

**Theorem 8.2**

Let  $A$  be an  $m \times n$  matrix. Then the following statements are equivalent:

- (a) For each  $\mathbf{b} \in \mathbb{R}^m$ , the equation  $A\mathbf{x} = \mathbf{b}$  has a solution.
- (b) Each  $\mathbf{b} \in \mathbb{R}^m$  is a linear combination of the columns of  $A$ .
- (c) The columns of  $A$  span  $\mathbb{R}^m$ .
- (d) The matrix  $A$  has a pivot position in every row.

**Row-Vector Rule for Computing  $A\mathbf{x}$** 

Assuming the product  $A\mathbf{x}$  is defined, the  $i$ -th entry in  $A\mathbf{x}$  is the sum of the products of corresponding entries from row  $i$  of  $A$  and the vector  $\mathbf{x}$ . In other words, the  $i$ -th entry of  $A\mathbf{x}$  is the dot product of the vector forming the  $i$ -th row of  $A$  and the vector  $\mathbf{x}$ .

**Theorem 8.3**

If  $A$  is an  $m \times n$  matrix,  $\mathbf{u}$  and  $\mathbf{v}$  are vectors in  $\mathbb{R}^n$ , and  $c$  is a scalar, then:

- (a)  $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$
- (b)  $A(c\mathbf{u}) = c(A\mathbf{u})$




$$\sim \begin{bmatrix} 1 & 0 & -\frac{2}{3} & 0 \\ 0 & 1 & -\frac{5}{3} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

This gives us the equations

$$\begin{aligned} x_1 - \frac{2}{3}x_3 &= 0 \implies x_1 = \frac{2}{3}x_3 \\ x_2 - \frac{5}{3}x_3 &= 0 \implies x_2 = \frac{5}{3}x_3 \end{aligned}$$

In vector form, the solution  $\mathbf{x}$  of the equation  $A\mathbf{x} = \mathbf{0}$  is written

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{2}{3}x_3 \\ \frac{5}{3}x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} \frac{2}{3} \\ \frac{5}{3} \\ 1 \end{bmatrix}$$

Since  $x_3$  can be anything, in geometric terms this solution set describes the line in  $\mathbb{R}^3$  extending from the origin through the point  $(\frac{2}{3}, \frac{5}{3}, 1)$ . 


### Parametric Vector Form

Whenever a solution set is described explicitly with vectors (as in the preceding example), we say that the solution is in **parametric vector form**.

### Theorem 8.4

Suppose  $A\mathbf{x} = \mathbf{b}$  is consistent for some  $\mathbf{b}$ , and let  $\mathbf{p}$  be a solution. Then the solution set of  $A\mathbf{x} = \mathbf{b}$  is the set of all vectors of the form

$$\mathbf{w} = \mathbf{p} + \mathbf{v}_h,$$

where  $\mathbf{v}_h$  is any solution of the homogeneous equation  $A\mathbf{x} = \mathbf{0}$ . 

### Definition 8.5

An indexed set of vectors  $\{\mathbf{v}_1, \dots, \mathbf{v}_p\} \subset \mathbb{R}^n$  is said to be **linearly independent** if the vector equation


$$x_1\mathbf{v}_1 + \dots + x_p\mathbf{v}_p = \mathbf{0}$$

has only the trivial solution, i.e., if the only solution is  $(x_1, \dots, x_p) = (0, \dots, 0)$ . Likewise, the set  $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$  is said to be **linearly dependent** if there exist weights  $c_1, \dots, c_p$ , not all zero, such that

$$c_1\mathbf{v}_1 + \dots + c_p\mathbf{v}_p = \mathbf{0}$$

We call such an equation a *linear dependence relation* when the weights are not all zero. 

### Theorem 8.5

- (a) If a set of vectors contains the zero vector, then the set is linearly dependent.
  - (b) If a set of vectors contains a scalar multiple of another vector, then the set is linearly dependent.
  - (c) If a set of vectors contains more vectors than there are entries in each vector, then the set is linearly dependent.
- 

### Corollary 8.1

An indexed set  $S = \{v_1, \dots, v_p\}$  of two or more vectors is linearly dependent if and only if at least one of the vectors in  $S$  is a linear combination of the others. If  $S$  is linearly dependent and  $v_1 \neq 0$ , then some  $v_j$  (with  $1 < j \leq p$ ) is a linear combination of the preceding vectors  $v_1, \dots, v_{j-1}$ .



## Matrix Algebra

### Definition 9.1 (Matrix Addition)

Let  $A$  and  $B$  be  $m \times n$  matrices. The **sum** of  $A$  and  $B$ , denoted  $A + B$ , is the  $m \times n$  matrix whose entries are obtained by adding the corresponding entries of  $A$  and  $B$ .

Given matrices

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

both of the same dimension  $m \times n$ , the sum  $A + B$  is thus defined as

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$



### Definition 9.2 (Scalar-Matrix Multiplication)

Let  $A$  be an  $m \times n$  matrix and  $c$  be a scalar. The **product** of  $c$  and  $A$ , denoted  $cA$ , is the  $m \times n$  matrix whose entries are obtained by multiplying each entry of  $A$  by  $c$ , and is thus defined as

$$cA = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}$$



### Theorem 9.1

Let  $A, B, C$  be matrices of the same size and let  $\alpha, \beta$  be scalars. Then

- (a)  $A + B = B + A$
- (b)  $(A + B) + C = A + (B + C)$
- (c)  $A + 0 = A$
- (d)  $\alpha(A + B) = \alpha A + \alpha B$
- (e)  $(\alpha + \beta)A = \alpha A + \beta A$
- (f)  $\alpha(\beta A) = (\alpha\beta)A$



**Definition 9.3**

For  $A \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{n \times p}$ , with  $B = \begin{bmatrix} b_1 & b_2 & \cdots & b_p \end{bmatrix}$ , we define the product  $AB$  by the formula

$$AB = \begin{bmatrix} Ab_1 & Ab_2 & \cdots & Ab_p \end{bmatrix}$$

**Theorem 9.2**

Let  $A, B, C$  be matrices, of appropriate dimensions, and let  $\alpha$  be a scalar. Then

- (a)  $A(BC) = (AB)C$
- (b)  $A(B + C) = AB + AC$
- (c)  $(B + C)A = BA + CA$
- (d)  $\alpha(AB) = (\alpha A)B = A(\alpha B)$
- (e)  $I_n A = A I_n = A$

**Definition 9.4**

Let  $A$  be a square matrix, i.e.  $A \in \mathbb{R}^{n \times n}$ . The  $k$ th power of  $A$ , denoted  $A^k$ , is defined as the product of  $A$  with itself  $k$  times. That is,

$$A^k = \underbrace{AAA \cdots A}_{k \text{ times}}$$

where  $A$  appears  $k$  times on the right-hand side.

**Definition 9.5**

Given a matrix  $A \in \mathbb{R}^{m \times n}$ , the transpose of  $A$  is the matrix  $A^T$  whose  $i$ th column is the  $i$ th row of  $A$ .

**Theorem 9.3**

Let  $A$  and  $B$  be matrices of appropriate dimensions and let  $\alpha$  be a scalar. Then

- (a)  $(A^T)^T = A$
- (b)  $(A + B)^T = A^T + B^T$
- (c)  $(\alpha A)^T = \alpha A^T$
- (d)  $(AB)^T = B^T A^T$

**Definition 9.6**

A square matrix  $A \in \mathbb{R}^{n \times n}$  is invertible (or **nonsingular**) if there exists a matrix  $C \in \mathbb{R}^{n \times n}$  such that  $AC = CA = I_n$ . The matrix  $C$  is called the inverse of  $A$  and is denoted by  $A^{-1}$ . Thus,  $A^{-1}A = AA^{-1} = I_n$ .





**Theorem 9.4**

Let  $A$  and  $B$  be invertible  $n \times n$  matrices. Then:

(a)  $A^{-1}$  is invertible, with

$$(A^{-1})^{-1} = A$$

(b) The product  $AB$  is invertible, with

$$(AB)^{-1} = B^{-1}A^{-1}$$

(c) The transpose of  $A$  is also invertible, i.e.  $A^T$  is invertible, with

$$(A^T)^{-1} = (A^{-1})^T$$

**Theorem 9.5**

Let  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ . If  $ad - bc \neq 0$ , then the inverse of  $A$  is given by

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$



**Remark:** The quantity  $ad - bc$  is called the determinant of  $A$ , and we write

$$\det(A) = |A| = ad - bc$$

**Theorem 9.6**

An  $n \times n$  matrix  $A$  is invertible if and only if  $A$  is row equivalent to  $I_n$ . In this case, any sequence of elementary row operations that reduces  $A$  to  $I_n$  also transforms  $I_n$  into  $A^{-1}$ .

**Algorithm for Finding the Inverse of an  $n \times n$  Matrix**

Let  $A$  be an  $n \times n$  matrix. To find the inverse of  $A$ , follow these steps:

1. Form the augmented matrix  $[A|I_n]$ .
2. Perform row operations on  $[A|I_n]$  to reduce  $A$  to  $I_n$ .
3. The matrix on the right side of the augmented matrix is  $A^{-1}$ .

If  $A$  is not invertible, then the algorithm will not be able to reduce  $A$  to  $I_n$ .

**Theorem 9.7**

Let  $A$  be an invertible matrix. Then the equation  $A\mathbf{x} = \mathbf{b}$  has a unique solution given by  $\mathbf{x} = A^{-1}\mathbf{b}$ .



**Theorem 9.8**

Let  $A$  be an  $n \times n$  matrix. The following statements are equivalent:

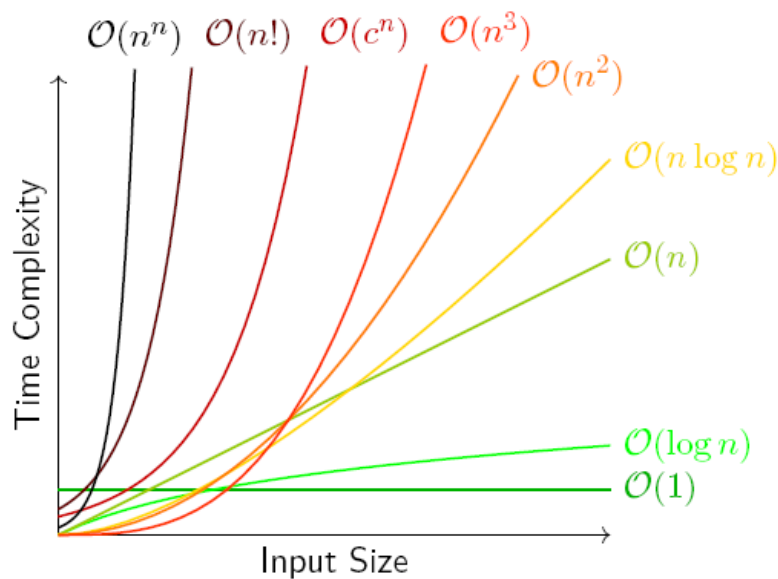
- (a)  $A$  is invertible.
- (b)  $A$  is row equivalent to  $I_n$ .
- (c)  $A$  has  $n$  pivot positions (i.e. one for each row and column).
- (d) The equation  $A\mathbf{x} = \mathbf{0}$  has only the trivial solution.
- (e) The columns of  $A$  are linearly independent.
- (f) The equation  $A\mathbf{x} = \mathbf{b}$  has a unique solution for each  $\mathbf{b} \in \mathbb{R}^n$ .
- (g) The columns of  $A$  span  $\mathbb{R}^n$ .
- (h)  $\det(A) \neq 0$ .
- (i) There is an  $n \times n$  matrix  $C$  such that  $CA = I$ .
- (j) There is an  $n \times n$  matrix  $D$  such that  $AD = I$ .
- (k)  $A^T$  is invertible.



## Time Complexity

Complexity	Growth Rate	Example Algorithms
$\mathcal{O}(1)$	Constant	Accessing an array element
$\mathcal{O}(\log n)$	Logarithmic	Binary search
$\mathcal{O}(n)$	Linear	Linear search
$\mathcal{O}(n \log n)$	Linearithmic (or Log-Linear)	MergeSort, QuickSort, HeapSort
$\mathcal{O}(n^2)$	Quadratic	Bubble Sort, Selection Sort
$\mathcal{O}(n^3)$	Cubic	Matrix multiplication
$\mathcal{O}(n^k)$	Polynomial	Polynomial-time approximation algorithms
$\mathcal{O}(c^n)$	Exponential	Subset sum, brute force for combinatorial problems
$\mathcal{O}(n!)$	Factorial	Brute force for travelling salesperson
$\mathcal{O}(n^n)$	Super-exponential	Recursive algorithms with high branching factor

**Table 10.1:** Growth rates of common functions and example algorithms



**Figure 10.3:** Conceptual illustration of time complexities

#### Definition 10.1 (Theta( $\Theta$ ))

$f(n)$  is  $\Theta(g(n))$  if there exist positive constants  $c_1, c_2$ , and  $k$  such that for all  $n \geq k$ ,

$$c_1 \cdot g(n) \leq f(n) \leq c_2 \cdot g(n)$$

This means that  $f(n)$  is asymptotically bounded both above and below by  $g(n)$ , meaning that  $f(n)$  grows at the same rate as  $g(n)$  as  $n$  becomes sufficiently large.



#### Proposition 10.1

Let  $f(n)$  and  $g(n)$  be two nonnegative functions, and suppose there is a constant  $c > 0$  for which

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = c$$

Then  $f(n) = \Theta(g(n))$ . Also, if

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)}$$

exists but does **not** equal a positive constant, then  $f(n) \neq \Theta(g(n))$ .



#### Definition 10.2 (Big- $\mathcal{O}$ )

$f(n)$  is  $\mathcal{O}(g(n))$  if there exist positive constants  $c$  and  $k$  such that for all  $n \geq k$ ,

$$f(n) \leq c \cdot g(n)$$


This means that  $f(n)$  is asymptotically bounded above by  $g(n)$ , meaning that  $f(n)$  grows at most as fast as  $g(n)$  as  $n$  becomes sufficiently large.



**Definition 10.3 (Omega Notation)**

$f(n)$  is  $\Omega(g(n))$  if there exist positive constants  $c$  and  $k$  such that for all  $n \geq k$ ,


$$f(n) \geq c \cdot g(n)$$

This means that  $f(n)$  is asymptotically bounded below by  $g(n)$ , meaning that  $f(n)$  grows at least as fast as  $g(n)$  as  $n$  becomes sufficiently large. 

**Definition 10.4 (Little-o)**

$f(n)$  is  $o(g(n))$  if for any positive constant  $\epsilon > 0$ , there exists a constant  $k \geq 0$  such that for all  $n \geq k$ ,

$$f(n) < \epsilon \cdot g(n)$$

This means that  $f(n)$  grows strictly slower than  $g(n)$  as  $n$  becomes sufficiently large. 

**Proposition 10.2**

Let  $f(n)$  and  $g(n)$  be two nonnegative functions, and suppose


$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 0$$

Then  $f(n) = o(g(n))$ , meaning  $f(n)$  grows strictly slower than  $g(n)$  as  $n \rightarrow \infty$ . 

**Definition 10.5 (Little- $\omega$ )**

$f(n)$  is  $\omega(g(n))$  if for any positive constant  $\epsilon > 0$ , there exists a constant  $k \geq 0$  such that for all  $n \geq k$ ,

$$f(n) > \epsilon \cdot g(n)$$

This means that  $f(n)$  grows strictly faster than  $g(n)$  as  $n$  becomes sufficiently large. 

**Proposition 10.3**

Let  $f(n)$  and  $g(n)$  be two nonnegative functions, and suppose

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = \infty$$

Then  $f(n) = \omega(g(n))$ , meaning  $f(n)$  grows strictly faster than  $g(n)$  as  $n \rightarrow \infty$ . 

**Informal Limit Approach**

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} \neq \infty \Rightarrow f(n) = \mathcal{O}(g(n))$$

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} \neq 0 \Rightarrow f(n) = \Omega(g(n))$$

**Note:** This approach is an informal guideline and not a strict mathematical definition. There are cases where this limit-based method may not apply, but it often provides a convenient tool for determining asymptotic behavior.

## Basic Operations

The time required by a function or procedure is proportional to the number of "basic operations" that it performs. Examples of basic operations include:

1. **Arithmetic Operation:** A single arithmetic operation, such as addition (+), multiplication (\*), subtraction (-), or division (/).
2. **Assignment:** Assigning a value to a variable, e.g.,  $x = 0$ .
3. **Test/Comparison:** Evaluating a condition, such as  $x == 0$  or  $i \leq n$ .
4. **Function Call:** Invoking a function or procedure, e.g., calling FIND-MAX-EVEN-SUM.
5. **Return Statement:** Returning a value from a function, e.g., `return total`.
6. **Read Operation:** Reading a value of a primitive type (e.g., integer, float, character, boolean) from input or memory.
7. **Write Operation:** Writing a value of a primitive type to output or memory.

These basic operations form the foundation of more complex operations, and each adds one unit time unit to the construct they are part of, i.e.  $O(1)$  time complexity. By counting the number of basic operations within a construct, we can estimate the time complexity of that construct. We will **only include 1-5** in our analysis and omit read and write operations as they are not as common in algorithm analysis.

Sum	Closed Form
$\sum_{k=0}^n ar^k \ (r \neq 1)$	$\frac{ar^{n+1} - a}{r - 1}$
$\sum_{k=1}^n k$	$\frac{n(n+1)}{2}$
$\sum_{k=1}^n k^2$	$\frac{n(n+1)(2n+1)}{6}$
$\sum_{k=1}^n k^3$	$\frac{n^2(n+1)^2}{4}$
$\sum_{k=0}^n x^k \ (x \neq 1)$	$\frac{x^{n+1} - 1}{x - 1}$
$\sum_{i=1,2,4,\dots,n} i$	$2n - 1$
$\sum_{k=0}^{\infty} x^k,  x  < 1$	$\frac{1}{1 - x}$
$\sum_{k=1}^{\infty} kx^{k-1},  x  < 1$	$\frac{1}{(1 - x)^2}$
$\sum_{k=1}^n \frac{1}{k}$	$\ln n + O(1)$
$\sum_{k=1}^n (a + bk)$	$an + b\frac{n(n+1)}{2}$

**Table A.1:** Some Useful Summation Formulae

The summation

$$\sum_{k=1}^n k = 1 + 2 + \cdots + n$$

is an arithmetic series and has the value

$$\begin{aligned}\sum_{k=1}^n k &= \frac{n(n+1)}{2} \\ &= \Theta(n^2).\end{aligned}$$

The following formulas apply to summations of squares and cubes:

$$\sum_{k=0}^n k^2 = \frac{n(n+1)(2n+1)}{6},$$

$$\sum_{k=0}^n k^3 = \frac{n^2(n+1)^2}{4}.$$

For real  $x \neq 1$ , the summation

$$\sum_{k=0}^n x^k = 1 + x + x^2 + \cdots + x^n$$

is a geometric series and has the value

$$\sum_{k=0}^n x^k = \frac{x^{n+1} - 1}{x - 1}.$$

The infinite decreasing geometric series occurs when the summation is infinite and  $|x| < 1$ :

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1 - x}.$$

If we assume that  $0^0 = 1$ , these formulas apply even when  $x = 0$ .

For positive integers  $n$ , the  $n$ th harmonic number is

$$\begin{aligned}H_n &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n} \\ &= \sum_{k=1}^n \frac{1}{k} \\ &= \ln n + O(1).\end{aligned}$$

The finite product  $a_1 a_2 \dots a_n$  can be expressed as

$$\prod_{k=1}^n a_k.$$

If  $n = 0$ , the value of the product is defined to be 1. You can convert a formula with a product to a formula with a summation by using the identity

$$\log \left( \prod_{k=1}^n a_k \right) = \sum_{k=1}^n \log a_k.$$

